International Journal of Science Engineering and Advance Technology, IJSEAT, Vol 2, Issue 12, December - 2014 | ISSN 2321-6905

International Journal of Science Engineering and Advance Technology

Maintain Data Integrity And Protection Of Private Label Information In Social Network Data

[1]Mallam Ujwala, [2]P.Ravi
[1]Mail id: ujwala566@gmail.com
1,2 Balaji Institute of Technology & Science Narsampet warangal

Abstract:
Perceptive information about users of the social networks should be protected. The confront is to plan methods to publish social network data in a form that affords usefulness without compromising privacy. Previous research has proposed a variety of privacy models with the corresponding protection mechanisms that put off both unintentional private information escape and attacks by malicious adversaries. These early privacy models are mainly disturbed with identity and link revelation. The social networks are modelled as graphs in which users are nodes and social connections are edges. The intimidation definitions and defence mechanisms leverage structural properties of the graph. This paper is stimulated by the recognition of the need for a better grain and more personalized privacy. Users commend social networks such as Face book and LinkedIn with a wealth of personal information such as their age, address, current location or political orientation. We refer to these details and messages as features in the user's profile. We propose a privacy protection scheme that not only put off the revelation of identity of users but also the disclosure of selected features in users' profiles.

Introduction:
The privacy problem happens from the revelation of sensitive labels. One might recommend that such labels should be simply deleted. Still such a solution would present an unfinished view of the network and may hide interesting statistical information that does not intimidate privacy. A more complicated approach consists in releasing information about sensitive labels while make certain that the identities of users are protected from privacy threats. We consider such threats as neighbourhood attack in which an adversary finds out responsive information based on prior knowledge of the number of neighbours of a target node and the labels of these neighbours. We recommend a privacy protection scheme that not only prevents the disclosure of identity of users but also the disclosure of selected features in users' profiles. An individual user can choose which features of her profile she wishes to hide. The social networks are modelled as graphs in which users are nodes and features are labels. Labels are denoted either as sensitive or as non-sensitive. We treat node labels both as background knowledge an adversary may possess and as sensitive information that has to be protected. We present privacy protection algorithms that consent to for graph data to be published in a form such that an adversary who possesses information about a node's neighbourhood cannot securely infer its identity and its sensitive labels. To this aspire the algorithms transform the original graph into a graph in which nodes are adequately impossible to differentiate. The algorithms are designed to do so while losing as small information and while protecting as much utility as possible.

Related Work:
The adolescent technique has quickly been recognized as deteriorating to protect privacy. For micro data Sweeney et al. propose k-anonymity to get round possible individuality disclosure in naively anonymized micro data. Diversity is proposed in order to additional prevent attribute disclosure. Similarly for network data Backstrom et al show that naive anonymization is insufficient as the arrangement of the released graph may expose the identity of the individuals corresponding to the nodes. Hay et al. highlight this problem and compute the risk of reidentification by adversaries with external information that is dignified into structural queries node refinement queries, sub graph knowledge queries. Be familiar with the problem several works propose procedure that can be applied to the naive anonymized graph further change the graph in order to afford certain privacy assurance. Some works are based on graph models other than simple graph. Zhou, Pei and Yuan et al. were the first to think about modelling social networks as labelled graphs similarly to what we think about in this paper. To prevent reidentification attacks by adversaries with immediate neighbourhood structural knowledge. They put into effect a k-anonymity privacy restraint on the graph each node of which is guaranteed to

have the same immediate neighbourhood structure with other k-1 nodes.

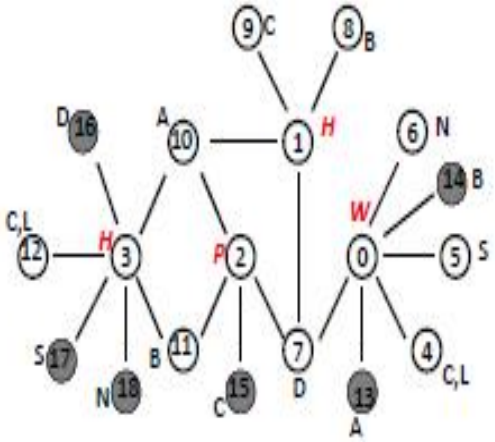
Existing System:
The present inclination in the Social Network it not generous the privacy about user profile views. The method of data sharing or Posting has taking additional time and not under the sure condition of displaying sensitive and non-sensitive data.

Disadvantages:
There is no way to publish the Non sensitive data to all in social Network. It's not providing privacy about user profiles. Some mechanisms that prevent both inadvertent private information leakage and attacks by malicious adversaries.

Proposed System:
We extend the existing definitions of modules and we introduced the sensitive or non-sensitive label concept in our project. We overcome the existing system disadvantages in our project.

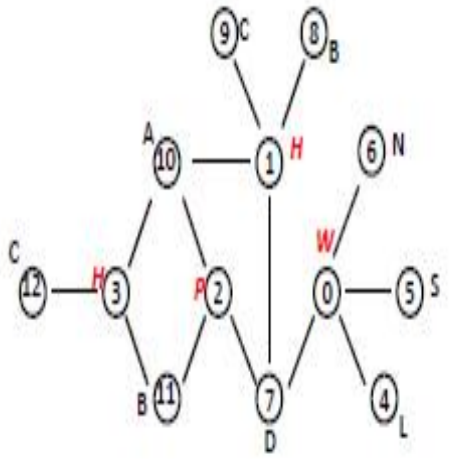
Advantages:
We can publish the Non sensitive data to every-one in social Network. It's providing privacy for the user profiles so that unwanted persons not able to view your profiles. We can post sensitive data to particular peoples and same way we can post non-sensitive data to everyone like ads or job posts.

Privacy Attaining Network:



We indulgence node labels both as part of an adversary's background knowledge and as sensitive information that has to be protected. The neighbourhood information of node v consists of the degree of v and the labels of v's neighbours. For each node v that connections with a sensitive label there must be at least. Other nodes with the same neighbourhood information but attached with different sensitive labels. Satisfies 2-sensitive-label-diversity that is as in this graph nodes 0 and 3 are indistinguishable having six neighbours with label A, B, fC, Lg, D, S, N separately similarly nodes 1 and 2 are indistinguishable as they both have four neighbours with labels A, B, C, D separately.

Labelled Graph Representation:



The social networks are modelled as graphs in which users are nodes and features are labels. Labels are designated either as sensitive or as non-sensitive. The above figure is a labelled graph representing a small subset of such a social network. Each node in the graph represents a user and the edge between two nodes represents the fact that the two persons are friends. Labels explained to the nodes show the locations of users. Each letter represents a city name as a label for each node. Some individuals do not mind their residence being known by the others but some do for a variety of reasons. In such case the privacy of their labels should be protected at data release. Therefore the locations are either sensitive or non-sensitive.

User Module:
Users are having confirmation and safety to access the detail which is presented in the ontology system. Before accessing or penetrating the details user should have the account in that otherwise they should register first.

Information Loss:
We aspire to remain information loss low. Information loss in this case encloses both structure information loss and label information loss. There are some non sensitive data's are loss due to Privacy making so we can't send out full information to the public.

Sensitive Label Privacy Protection:
There is who post the image to the online social network if permit the people for showing the image it will exhibit to his requesters it make as the sensitive to that user. This is very helpful to make sensitive data for the public.

Algorithm Used:

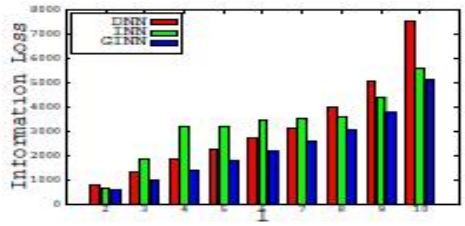
Algorithm 1: Global-Similarity-based Indirect Noisy Node Algorithm

Input: graph $G(V, E, L, L')$, parameter l;
Result: Modified Graph G'

```
while V_left > 0 do
    if |V_left| ≥ l then
        compute pairwise node similarities;
        group G ← v1, v2 with Max_similarity;
        Modify neighbors of G;
        while |G| < l do
            dissimilarity(V_left, G);
            group G ← v with Max_similarity;
            Modify neighbors of G without actually adding noisy nodes ;
    else if |V_left| < l then
        for each v ∈ V_left do
            similarity(v, Gs);
            G_Max_similarity ← v;
        Modify neighbors of G_Max_similarity without actually adding noisy
        nodes;
Add expected noisy nodes;
Return G'(V', E', L');
```

Noise node addition process that is accepted to make the nodes inside each group satisfies sensitive-label-diversity is documented but not performed right away. Only after all the initial grouping operations are performed the algorithm precedes to procedure the expected node addition operation at the final step. Then if two nodes are expected to have the same labels of neighbours and are within two hops having common neighbours only one node is added. In other words we combine some noisy nodes with the same label thus resulting in fewer noisy nodes.

Experimental Results:



We evaluate the data utilities we conserve from the original graphs in view of measurements on degree distribution, label distribution, degree centrality, clustering coefficient, average path length, graph density and radius. We show the numeral of the noisy nodes and edges needed for each approach. In view of utility of released data we intend to keep information loss low. Information loss in this case contains both structure information loss and label information loss. The measurements of information loss on the artificial data set using each algorithm. Algorithm GINN introduces the least information loss.

Conclusion:

We suppose that adversaries possess previous knowledge about a node's degree and the labels of its neighbours and can use that to deduce the sensitive labels of targets. We recommended a model for accomplish privacy while publishing the data in which node labels are both part of adversaries' background knowledge and susceptible information that has to be protected. We go together with our model with algorithms that change a network graph before publication so as to bound adversaries' confidence about sensitive label data. Our experiments on both real and synthetic data sets confirm the helpfulness, competence and scalability of our approach in maintaining essential graph properties while providing a understandable privacy guarantee.

References:

1. L. A. Adamic and N. Glance. The political blogosphere and the 2004 U.S. election: divided they blog. In LinkKDD, 2005.
2. L. Backstrom, C. Dwork, and J. M. Kleinberg. Wherefore art thou R3579X?: anonymized social networks, hidden patterns, and structural steganography. Com- mun. ACM, 54(12), 2011.
3. S. Bhagat, G. Cormode, B. Krishnamurthy, and D. S. and. Class-based graph anonymization for social network data. PVLDB, 2(1), 2009.
4. A. Campan and T. M. Truta. A clustering approach for data and structural anonymity in social networks. In PinKDD, 2008.
5. J. Cheng, A. W.-C. Fu, and J. Liu. K-isomorphism: privacy-preserving network publication against structural attacks. In SIGMOD, 2010.
6. G. Cormode, D. Srivastava, T. Yu, and Q. Zhang. Anonymizing bipartite graph data using safe groupings. PVLDB, 19(1), 2010.
7. S. Das, O. Egecioglu, and A. E. Abbadi. Anonymizing weighted social network graphs. In ICDE, 2010.
8. A. G. Francesco Bonchi and T. Tassa. Identity obfuscation in graphs through the information theoretic lens. In ICDE, 2011.
9. M. Hay, G. Miklau, D. Jensen, D. Towsley, and P. Weis. Resisting structural re-identi_cation in anonymized social networks. PVLDB, 1(1), 2008.
10. Y. Li and H. Shen. Anonymizing graphs against weight-based attacks. In ICDM Workshops, 2010.
11. K. Liu and E. Terzi. Towards identity anonymization on graphs. In SIGMOD, 2008.
12. L. Liu, J.Wang, J. Liu, and J. Zhang. Privacy preserving in social networks against sensitive edge disclosure. In SIAM International Conference on Data Mining, 2009.
13. A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam. `-diversity: privacy beyond k-anonymity. In ICDE, 2006.